

TOEIC Speaking Test and Willingness to Communicate

Masaya Kanzaki

Kanda University of International Studies

Reference Data:

Kanzaki, M. (2016). TOEIC Speaking test and willingness to communicate. In P. Clements, A. Krause, & H. Brown (Eds.), *Focus on the learner*. Tokyo: JALT.

In this study was explored the relationship between university students' scores on the TOEIC Speaking test and levels of willingness to communicate (WTC) as well as their scores on the TOEIC Listening and Reading test. While some learners are able to achieve high scores on the speaking test despite having relatively low listening and reading test scores, others receive low speaking test scores despite their high listening and reading test scores, and WTC is considered to be a factor affecting their speaking test scores. In this study, the TOEIC Speaking test, the TOEIC Listening and Reading Test, and a 9-item WTC questionnaire were given to 142 university students. Correlations between the scores of the tests and the questionnaire responses were then examined and a multiple regression analysis was performed. The results indicated a significant relationship between speaking test scores and levels of WTC. In addition, the pedagogical and practical implications of the study are discussed.

本研究は、TOEICスピーキングテストの得点とwillingness to communicate (WTC:話をしたいという気持ち)の程度、並びにTOEICリスニング・リーディングテストのスコアとの関係を探った。リスニングとリーディングの得点が低いにも関わらず、スピーキングテストで高得点を取ることができる学習者がいる一方で、高いリスニングとリーディングの得点にも関わらずスピーキングテストのスコアが低い者もあり、WTCは学習者のスピーキングテストの得点に影響を与えるひとつの要因として考えられている。本研究では、TOEICスピーキングテスト、TOEICリスニング・リーディングテスト、そして9項目のWTC調査票が142人の大学生に対して与えられた。そして、テストの得点と調査票への回答の間の相関関係が分析され、重回帰分析が行われた。研究結果はスピーキングテストの得点とWTCレベルの間に有意な関係があることを示した。さらに本研究の教育上および実用面での意義についても言及する。

Educational Testing Service (ETS), the developer of the TOEIC Test, launched the TOEIC Speaking test, a computer-based speaking test, in 2006 to complement the

all multiple-choice TOEIC Listening and Reading test. The speaking test is designed to measure learners' abilities "to communicate in spoken English in the context of daily life and the global workplace" (ETS, 2010, p. 6). A few ETS-commissioned correlation studies have compared speaking, listening, and reading scores on the TOEIC test. Liao, Qu, and Morgan (2010) reported correlations of .66 between the speaking and listening scores and .57 between the speaking and reading scores ($N = 12,099$). Liu and Costanzo (2013) found correlations of .63 between the speaking and listening scores and .54 between the speaking and reading scores ($N = 4,935$). While statistically significant, these correlations were not particularly high, which means there were test takers who had high speaking scores and low listening scores, reading scores, or both, and others who had low speaking scores and high listening scores, reading scores, or both. Why do some learners speak well despite their weak receptive skills while other learners with strong receptive skills speak poorly? One possible answer is their different levels of willingness to communicate (WTC). It is conceivable that those who are more willing to communicate verbally with others are better at speaking than those who are less willing to do so.

MacIntyre, Clément, Dörnyei, and Noels (1998) defined WTC as "a readiness to enter into discourse at a particular time with a specific person or persons, using a[n] L2" (p. 547), and this definition has been widely quoted in EFL literature. However, WTC research originated from behavioral studies into the differences between individuals in L1 communication, which investigated the causes of their unwillingness to communicate (e.g., Burgoon, 1976; McCroskey & Richmond, 1982; Mortensen, Arntson, & Lustig, 1977). McCroskey and Richmond (1990) argued that whether a person is willing to communicate or not is a cognitively processed volitional choice and that "the personality of the individual might be the determining factor" of that choice (p. 21), implying that WTC in one's first language is a personal trait. In L2 communication, however, situational variables, such as intergroup issues, influence WTC as well as trait-like variables. MacIntyre et al. (1998) organized variables influencing WTC in L2 into six layers: communication behavior, behavioral intention, situated antecedents, motivational propensities, affec-

Kanzaki: TOEIC Speaking Test and Willingness to Communicate

tive-cognitive context, and social and individual context and 12 subdivisions—including self-confidence, intergroup motivation, and communication competence—in a pyramid model. In L2 WTC research, questionnaires are often administered to measure levels of WTC (e.g., Doe, 2014; MacIntyre, Baker, Clément, & Conrod, 2001; Ockert, 2012; Weaver, 2010; Yashima, Zenuk-Nishide, & Shimizu, 2004).

There have been a few studies comparing the results of English language tests and levels of WTC. Two such studies were conducted in Iran recently but offered contradictory results. Baghaei and Dourakhshan (2012) compared the results of a 20-item WTC questionnaire with those of a 20-item cloze test and found a moderate correlation between them, $r = .39, p < .01$ (two-tailed), $N = 148$. Rahmatollahi and Khalili (2015) later compared the results of a 20-item WTC questionnaire with those of an intermediate-level speaking test and found no significant relationship, $\rho = .029, p = .751$ (two-tailed), $N = 120$.

To my knowledge, no study has been conducted to examine the relationship between TOEIC scores and WTC, so carrying out such a study might be worthwhile. The research questions for this study are as follows:

- RQ1. Do the levels of WTC differ between people with high TOEIC speaking scores but low TOEIC listening and reading scores and those with low TOEIC speaking scores but high TOEIC listening and reading scores?
- RQ2. To what extent do TOEIC speaking, listening, and reading scores correlate with each other?
- RQ3. To what extent do levels of WTC correlate with TOEIC speaking, listening, and reading scores?
- RQ4. To what extent can levels of WTC predict TOEIC speaking scores?

Method

Data used in this study were collected in July 2014 and July 2015. The TOEIC Speaking test and the TOEIC Listening and Reading test as well as a questionnaire to determine levels of WTC in verbal communication in English (WTC questionnaire) were administered, the scores of which were then computed for statistical analyses.

Participants

The participants of this study were 142 university students who were attending a private Japanese university in the Kanto region that specializes in foreign languages. Each student was given 1,000 yen to take part in the study, and they each paid 3,500 yen to take the TOEIC Listening and Reading test. The cost of the TOEIC Speaking test was covered

by a research grant. In 2014, 94 students signed up for the study, however three of them were excluded because they did not complete the WTC questionnaire properly. In 2015, 54 students signed up, but three were excluded for the same reason. There were four students who participated in both years; however, their 2nd-year results were different from their results in the 1st year, and thus the data from both years were used in this study.

The purposes, procedures, and requirements of the study were explained to the participants before they signed a consent form. They understood that the TOEIC Speaking test and the questionnaire would not affect their grades in any way and that the results of the TOEIC Listening and Reading test would affect their grades in the TOEIC course.

Among the 142 participants, nine were in their 1st academic year, 70 in their 2nd year, 34 in their 3rd year, and 29 in their 4th year; 24 of the participants were male and 118 were female. In terms of fields of study, there were 77 international communication majors, 40 English language majors, 17 international business majors, three Portuguese language majors, two Indonesian language majors, one Chinese language major, one Spanish language major, and one Vietnamese language major. Four were enrolled in TOEIC-860 courses, six in TOEIC-730 courses, 59 in TOEIC-650 courses, and 10 in TOEIC-600 courses (860, 730, 650, and 600 indicate the targeted TOEIC scores of the courses). The remaining 63 were not enrolled in any TOEIC courses.

Materials and Procedures

A WTC questionnaire and the TOEIC speaking, listening, and reading tests were used in this study. The listening and reading tests are always administered together, and are therefore generally treated as two subtests of one test. The speaking test, on the other hand, can be taken independently or with the TOEIC Writing test. The TOEIC tests used in this study were administered in the Institutional Program (IP), for which an institution sets the time, date, and place of the exam. The questionnaire was administered via SurveyMonkey, an online survey administration service, and completed before the participants took the tests. The survey data were collected online and exported into a Microsoft Excel spreadsheet. The TOEIC tests were administered over 2 days; the TOEIC Listening and Reading test on the 1st day and the TOEIC Speaking test on the 2nd day. All participants took the TOEIC Listening and Reading test at the same time, however due to a lack of computers, the TOEIC Speaking test was taken at different times, either in the morning or in the afternoon. The results of the TOEIC tests were provided by the Institute for International Business Communication, the administrator of TOEIC in Japan. The scores were then analyzed using Microsoft Excel 2013, IBM SPSS Statistics 23, and IBM SPSS Bootstrapping 23.

WTC Questionnaire

A 9-item questionnaire was developed to determine the participants' levels of WTC in verbal communication in English. The questionnaire items were modeled after those in Weaver (2010) and Doe (2014). When developing their questionnaires, both Weaver and Doe chose situations and tasks that their students had frequently encountered in English classes and therefore would be able to imagine easily. The same principle was applied in the development of the questionnaire for this study. The items focused exclusively on verbal communication in the context of an English class. For the sake of simplicity and clarity, the Japanese phrases *yaru* (I do it) and *waranai* (I don't do it) were used in the answer choices, which concerned actual communication behavior rather than intention for actions. However, the Japanese phrase *donokurai sekkyokutekini* (how actively) was included in the directions so as to capture levels of willingness to take such actions. The questionnaire was piloted with 12 students; modifications were made based on their feedback. When administered to the 142 participants, the questionnaire proved to be fairly reliable, with a reliability index (Cronbach's alpha) of .81. For the questionnaire items and the 4-item Likert scale that was used, see Appendices A and B.

Three TOEIC Tests

The TOEIC Speaking test is a computer-based examination requiring test takers to sit in front of a computer while wearing a headset with a microphone. Instructions are provided on the computer screen and through the headset. Test takers speak into the microphone, and their speech is recorded and sent to certified raters for evaluation. There are 11 questions in the test, and scores are given in the range of 0 to 200.

The TOEIC Listening and Reading test is composed of two subtests. Each subtest consists of 100 multiple-choice questions, and raw scores of between 0 and 100 are converted to scaled scores of between 5 and 495. The listening test has four parts; the reading test has three. Details of the three tests are shown in Appendix C.

As ETS does not make item-by-item results of the TOEIC available, reliability indices for the three TOEIC tests used in this study could not be determined. However, ETS (2010) reported that the reliability of the TOEIC Speaking test was .80 based on the data from tests administered in the Public Testing Program from January 2008 to December 2009. Also, ETS (2013) reported that the reliability index (KR-20) of the TOEIC listening and reading scores across all forms of their norming samples was approximately .90.

Results and Discussion

Descriptive Statistics and Normality Tests

Table 1 shows the descriptive statistics for the scores of the three tests and the WTC questionnaire. Remarkably, the average listening score was 101.16 points higher than the average reading score.

In addition, histograms, normal Q-Q plots, and box plots for the scores were created in order to examine the normality of score distribution. A visual inspection of these confirmed that the scores of the speaking test, the reading test, the listening and reading tests combined, and the WTC questionnaire were approximately normally distributed. However, the distribution of listening scores was not normal, with a skewness of -0.73 ($SE = .20$) and a kurtosis of 1.15 ($SE = .40$), and therefore the assumption of normality was violated.

Table 1. Descriptive Statistics for the Three Tests and WTC Questionnaire ($N = 142$)

Test	Possible scores	Mean	SD	Min	Max	Skewness	Kurtosis
TS	0–200	117.61	21.63	60	180	0.22	0.04
TL	5–495	374.19	58.64	170	495	-0.73	1.15
TR	5–495	273.03	74.84	100	475	0.12	-0.56
TLR	10–990	647.23	122.23	310	945	-0.19	0.06
WTC	9–36	26.10	4.54	12	36	-0.10	-0.23

Note. TS = TOEIC Speaking test, TL = TOEIC Listening test, TR = TOEIC Reading test, TLR = TL and TR combined, WTC = WTC questionnaire, Min = lowest score, Max = highest score.

High Score Group Versus Low Score Group

In order to determine whether the levels of WTC differed between those with high and those with low TOEIC scores, the participants were divided into four groups corresponding to their speaking scores and combined listening and reading scores. The median for the speaking test was 120 and 52 and 65 participants scored above and below the medi-

Kanzaki: TOEIC Speaking Test and Willingness to Communicate

an, respectively. The top 52 participants were categorized as speaking high scorers (SH) and the bottom 65 as speaking low scorers (SL).

The median for the combined listening and reading scores was 645; one participant scored 650 and another scored 655. There were 68 participants who scored above 655 and another 68 who scored below 645. The top 68 participants were categorized as listening and reading high scorers (LRH) and the bottom 68 as listening and reading low scorers (LRL).

The participants were then put into the following four groups corresponding to the two categories in which they had been placed: SH-LRH (high speaking score with a high listening and reading score), SH-LRL (high speaking score with a low listening and reading score), SL-LRH (low speaking score with a high listening and reading score), and SL-LRL (low speaking score with a low listening and reading score). To make the differences between the high and low score groups clearer, those in the middle who scored either 120 on the speaking test or between 645 and 655 on the listening and reading tests, or both, were excluded. Table 2 shows the average WTC scores of the four groups.

The highest average WTC score among the four groups was found in the SH-LRL group ($M = 28.40$, $n = 11$) and the lowest in the SL-LRH group ($M = 24.00$, $n = 15$), which indicates that those with a low combined listening and reading score but who can speak well have a relatively high level of WTC, and those with a high combined listening and reading score but who cannot speak well have a relatively low level of WTC. There was only a small difference of .20 in the average WTC scores between the SH-LRH group and SH-LRL group. Similarly, the difference between the SL-LRH group and SL-LRL group was only 1.10. These results suggest that if their speaking scores were at similar levels, their WTC scores were also similar even when their combined listening and reading scores were at different levels. Conversely, the differences between the SH-LRH and SL-LRH groups and between the SH-LRL and SL-LRL groups were both relatively big, with the former being 4.20 and the latter 3.30. This suggests that if their speaking scores were at different levels, their WTC scores were different even when their combined listening and reading scores were at similar levels.

Table 2. Average WTC Scores for High and Low Scoring Groups ($N = 112$)

Speaking	LRH	LRL
SH	28.20 ($n = 41$)	28.40 ($n = 11$)
SL	24.00 ($n = 15$)	25.10 ($n = 45$)

Note. SH = high speaking score (> 120); SL = low speaking score (< 120); LRH = high listening and reading score (> 165); LRL = low listening and reading score (< 155).

Correlations

Scores for Speaking, Listening, and Reading Tests

Table 3 shows the correlations between the scores of the three tests. As the assumption of normality was violated, the bootstrap function of SPSS Statistics was activated when the correlations were calculated in order to overcome bias caused by nonnormality of score distribution (for more information about bootstrapping with SPSS, see Field, 2013). Bias corrected and accelerated bootstrap 95% confident intervals are reported in square brackets. Among the three combinations, the highest correlation was found between the listening and reading scores, $r = .67$ [.575, .749]. It is unusual that the speaking score is more closely correlated with the writing score, $r = .50$ [.342, .615], than the listening score, $r = .47$ [.344, .584] (all $ps < .001$). In other correlation studies involving these three tests, the correlation between the listening and speaking scores was usually higher than the correlation between the reading and speaking scores (e.g., Kanzaki, 2015; Liao et al., 2010; Liu & Constanzo, 2013). The lower correlation between the listening and speaking scores could be due to the nonnormality of listening score distribution. The speaking test score was also significantly correlated with the combined listening and reading score, $r = .53$ [.396, .644], $p < .001$ (two-tailed).

Kanzaki: TOEIC Speaking Test and Willingness to Communicate

Table 3. Correlations Between Speaking, Listening, and Reading Scores ($N = 142$)

Test	TS	TL	TR
TS	1	.47 [.344, .584]	.50 [.342, .615]
TL		1	.67 [.575, .749]
TR			1

Note. TS = TOEIC Speaking test; TL = TOEIC Listening test; TR = TOEIC Reading test; BCa bootstrap 95% CIs reported in brackets.

All $ps < .001$ (two-tailed).

WTC Versus TOEIC

Table 4 shows the correlations between the WTC scores and the TOEIC scores. As above, the bootstrap function of SPSS was activated when they were computed. The WTC score was significantly correlated with the speaking score, $r = .39$, 95% BCa CI [.254, .500], $p < .001$. However, the WTC score was not significantly correlated with the listening score, the reading score, or the combined listening and reading score, with rs being .12, .11, and .12, respectively, and ps being .149, .206, and .142, respectively.

Table 4. Correlations Between WTC and TOEIC ($N = 142$)

Test	TS	TL	TR	TS
	.39	.12	.11	.12
WTC	[.254, .500] $p < .001$	[-.049, .290] $p = .149$	[-.054, .263] $p = .206$	[-.048, .286] $p = .142$

Note. TS = TOEIC Speaking test; TL = TOEIC Listening test; TR = TOEIC Reading test; TLR = TL and TR combined; WTC = WTC questionnaire; BCa bootstrap 95% CIs reported in brackets.

All p values are two-tailed.

Multiple Regression Analysis

As shown in Table 4, the listening score, reading score, and WTC score were all moderately correlated with the speaking score and it was therefore assumed that they could be

predictors of the speaking score. In order to examine how well these three scores could predict the speaking score, a multiple regression analysis was conducted. A significant regression equation was found, $F(3, 138) = 29.01$, $p < .001$, with an R^2 of .39, which indicates that approximately 39% of the variance in the speaking score can be accounted for by the combination of the three predictors. The predicted speaking score of the participants was equal to $21.86 + 0.08$ (listening score) + 0.09 (reading score) + 1.56 (WTC score). The speaking score of the participants increased by 0.08 for each point on the listening test, 0.09 for each point on the reading test, and 1.56 for each point on the WTC questionnaire. The listening score, reading score, and WTC score were all significant predictors of the speaking score.

Table 5 summarizes unstandardized and standardized regression coefficients, standard errors for unstandardized coefficients, and p values for the linear model of predictors. The standardized beta values for the reading score and the WTC score were .32 and .33, respectively, indicating that both variables were equally important in the model. The standardized beta value for the listening score was .22, which is significantly lower than the other two scores.

Table 5. Linear Model of Predictors of Speaking Score ($N = 142$)

Score	B	$SE B$	β	p
Constant	21.86 [-1.00, 43.75]	10.91		$p = .050$
Listening	0.08 [0.01, 0.15]	0.03	.22	$p = .016$
Reading	0.09 [0.04, 0.14]	0.03	.32	$p = .003$
WTC	1.56 [0.97, 2.09]	0.28	.33	$p = .001$

Note. B = unstandardized regression coefficient; β = standardized regression coefficient; BCa bootstrap 95% CIs reported in brackets; Confidence intervals and standard errors are based on 1,000 bootstrap samples.

$R^2 = .39$ ($p < .001$). All p values are two-tailed.

Conclusions

The results presented indicate a clear relationship between the participants' speaking test scores and their levels of WTC. The average WTC scores were higher among the participants with high speaking scores compared to those with low speaking scores; the WTC score was significantly correlated with the speaking score; and the WTC score was a significant predictor of the speaking score. However, the findings do not indicate a causal relationship. Whether a change in WTC leads to a change in speaking test scores, or the other way around, is not clear. It may well be that levels of WTC are determined by levels of speaking ability, as some studies have suggested. For example, MacIntyre (1994) proposed that learners' perceived communication competence is responsible for determining their levels of WTC. Also, MacIntyre et al. (1998) pointed out that "one's degree of L2 proficiency will have a significant effect on his or her WTC" (p. 554). It seems natural that those with high levels of speaking ability are confident about their speaking skills and are therefore likely to have higher levels of WTC than those with lower communication competence. Even so, common sense suggests that WTC and speaking ability feed each other. As MacIntyre et al. (2001) remarked, an increase in WTC leads to more speaking in L2, which improves learners' speaking ability, and when their speaking ability improves, they become more willing to communicate. In this regard, an increase in WTC will ultimately lead to an improvement in speaking ability, and thus a pedagogical implication is that language teachers seeking to improve the speaking ability of their students should make an effort to foster WTC.

The following and final suggestion concerns the practical issue of testing. The results indicated that the WTC score could predict the speaking score better than the TOEIC Listening test and as well as the TOEIC Reading test. Although a speaking test can measure the speaking ability of students better than a paper test, it is generally more expensive and often logistically difficult to implement. Therefore, some institutions use the TOEIC Listening and Reading test as a placement test for speaking classes. In such cases, if they were to administer a simple WTC questionnaire with about a dozen items in addition to the TOEIC Listening and Reading test, they might be able to predict the speaking ability of students more accurately.

Acknowledgments

The author would like to thank Timothy Doe of Kwansei Gakuin University for sharing WTC-related materials, including his own questionnaire, without which this study would not have been possible. This study was supported by a JSPS KAKENHI Grant Number 25370727.

Bio Data

Masaya Kanzaki teaches at Kanda University of International Studies. His research interests include language testing, vocabulary acquisition, and corpus linguistics. <kanzaki-m@kanda.kuis.ac.jp>

References

- Baghaei, P., & Dourakhshan, A. (2012). The relationship between willingness to communicate and success in learning English as a foreign language. *Modern Journal of Applied Linguistics*, 4, 53-67.
- Burgoon, J. K. (1976). The unwillingness-to-communicate scale: Development and validation. *Communication Monographs*, 43(1), 60-69. <http://dx.doi.org/10.1080/03637757609375916>
- Doe, T. (2014). Willingness to communicate and group cohesion. In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings* (pp. 276-283). Tokyo: JALT.
- Educational Testing Service. (2010). *TOEIC user guide: Speaking and writing*. Princeton, NJ: Author. Retrieved from <http://www.ets.org/s/toEIC/pdf/toEIC_sw_score_user_guide.pdf>
- Educational Testing Service. (2013). *TOEIC user guide: Listening and reading*. Princeton, NJ: Author. Retrieved from <http://www.ets.org/Media/Tests/Test_of_English_for_International_Communication/TOEIC_User_Gd.pdf>
- Field, A. P. (2013). *Discovering statistics using IBM SPSS: And sex and drugs and rock'n'roll* (4th ed.). London: Sage.
- Kanzaki, M. (2015). TOEIC survey: Speaking vs. listening and reading. In P. Clements, A. Krause & H. Brown (Eds.), *JALT2014 Conference Proceedings* (pp. 639-649). Tokyo: JALT.
- Liao, C., Qu, Y., & Morgan, R. (2010). The relationships of test scores measured by the TOEIC listening and reading test and TOEIC speaking and writing tests. *TOEIC compendium 2*. Princeton, NJ: Educational Testing Service. Retrieved from <<https://www.ets.org/Media/Research/pdf/TC-10-13.pdf>>
- Liu, J., & Costanzo, K. (2013). The relationship among TOEIC listening, reading, speaking, and writing skills. *TOEIC compendium 2*. Princeton, NJ: Educational Testing Service. Retrieved from <<https://www.ets.org/Media/Research/pdf/TC2-02.pdf>>
- MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports*, 11(2), 135-142. <http://dx.doi.org/10.1080/08824099409359951>
- MacIntyre, P. D., Baker, S. C., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition*, 23, 369-388. <http://dx.doi.org/10.1017/S0272263101003035>
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, 82, 545-562. <http://dx.doi.org/10.1111/j.1540-4781.1998.tb05543.x>

Kanzaki: TOEIC Speaking Test and Willingness to Communicate

- McCroskey, J. C., & Richmond, V. P. (1982). Communication apprehension and shyness: Conceptual and operational distinctions. *Central States Speech Journal*, 33, 458-468. <http://dx.doi.org/10.1080/10510978209388452>
- McCroskey, J. C., & Richmond, V. P. (1990). Willingness to communicate: A cognitive view. *Journal of Social Behaviour and Personality*, 5(2), 19-37.
- Mortensen, C. D., Arntson, P. H., & Lustig, M. (1977). The measurement of verbal predispositions: Scale development and application. *Human Communication Research*, 3(2), 146-158. <http://dx.doi.org/10.1111/j.1468-2958.1977.tb00513.x>
- Ockert, D. (2012). Do Japanese JHS students have ideal L2 selves? Evidence from research in progress on the influence of multimedia use on affective variables. In A. Stewart & N. Sonda (Eds.), *JALT2011 Conference Proceedings* (pp. 168-178). Tokyo: JALT.
- Rahmatollahi, M., & Khalili, G. F. (2015). Relationship between intermediate EFL learners' communication apprehension, willingness to communicate, and speaking ability. *International Journal of Applied Linguistics and English Literature*, 4(6), 23-32. <http://dx.doi.org/10.7575/aiac.ijalel.v4n.6p.23>
- Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, 54, 119-152. <http://dx.doi.org/10.1111/j.1467-9922.2004.00250.x>
- Weaver, C. (2010). *Japanese university students' willingness to use English with different interlocutors* (Unpublished doctoral dissertation). Tokyo: Temple University Japan.

Appendix A

Willingness to Communicate Questionnaire Items (Japanese)

Items were answered on a 4-point Likert scale.

これまでの英語の授業の中で(授業での使用言語は英語)、次の1~9のことを英語でする場合、あなたはどのくらい積極的に行っていますか。それぞれの項目について自分の評価に近いものを4つの選択肢(1. 絶対やらない、2. あまりやらない、3. 多分やる、4. 確実にやる)からひとつ選んでください。

1. よく聞き取れなかったので、もう一度言ってもらうように頼む。
2. メモを見ながら自分のことについて短いスピーチをする。
3. みんなの前でロールプレイをする(例えば、ペアでレストランのウェイターと客の役になって注

文をする場面を演じる)。

4. 単語の発音の仕方を尋ねる。
5. 自分の経験したことについて話す。
6. 相手の経験について尋ねる。
7. 知らない単語の意味を尋ねる。
8. 何も見ずに自分のことについて短いスピーチをする。
9. ディスカッションで最初に発言する。

Appendix B

Willingness to Communicate Questionnaire (English Translation)

Items were answered on a 4-point Likert scale.

In English classes taught in English, how actively do you perform the following tasks in English? For each question, choose one of the following four responses that is closest to your experience: (1) I definitely don't do it, (2) I don't do it often, (3) I probably do it, and (4) I definitely do it.

1. Ask someone to repeat what they have just said because you did not catch it.
2. Give a short speech about yourself.
3. Do a role-play in front of the class (e.g., acting out a restaurant scene in which a customer orders food).
4. Ask someone how to pronounce a word.
5. Talk about your past experiences.
6. Ask someone about their past experiences.
7. Ask the meaning of a word you do not know.

Kanzaki: TOEIC Speaking Test and Willingness to Communicate

8. Give a short speech without notes.
9. Be the first person to speak in a discussion.

Appendix C
Details of the TOEIC Speaking, Listening, and Reading Tests

Table 1C. Details of the TOEIC Speaking Test

Question no.	Task
1, 2	Read aloud the text that appears on the screen.
3	Describe the picture on the screen.
4–6	Answer three questions about a single topic as though you are participating in a telephone interview.
7–9	Read the information on the screen and answer three questions about it as though you are responding to a telephone inquiry.
10	Listen to a recorded message about a problem and propose a solution for it.
11	Express an opinion about a specific topic.

Table 2C. Four Parts of the TOEIC Listening Test

Part	Task	No. of Qs
1	For each question with a photo, listen to four sentences and choose the one that best describes the image.	10
2	Listen to a question or statement followed by three responses and choose the most appropriate response.	30

3	Listen to a conversation and answer comprehension questions.	30
4	Listen to a short talk and answer comprehension questions.	30

Table 3C. Three Parts of the TOEIC Reading Test

Part	Task	No. of Qs
5	Choose a word or phrase to fill in a blank in a sentence.	40
6	Choose words or phrases to fill in blanks in a passage.	12
7	Read a passage or a set of two passages and answer comprehension questions.	48

Note. The TOEIC Reading test starts with Part 5 because it immediately follows the TOEIC Listening test, which ends with Part 4. The two tests are always taken as a set and often treated as two subtests of one test, the TOEIC Listening and Reading test.